The Effectiveness of Oral Resonance Therapy on the Perception of Femininity of Voice in Male-to-Female Transsexuals

Lisa Carew, Georgia Dacakis, and Jennifer Oates

Bundoora, Victoria, Australia

Summary: Ten male-to-female transsexuals participated in five sessions of oral resonance voice therapy targeting lip spreading and forward tongue carriage. Acoustic analysis of recordings made pre- and posttherapy found that participant formant frequency values (F1, F2, and F3, from the vowels /a/, /i/, and /ʊ/), as well as fundamental frequency (F0), underwent a general increase posttherapy. F3 values, in particular, increased significantly posttreatment. Trends in listener ratings of these recordings showed that the majority of participants were perceived to sound more feminine following treatment. Participants' self-ratings of their voices pre- and posttreatment also indicated that participants perceived their voices as sounding more feminine and that they were more satisfied with their voices following treatment. The present study supports the findings of previous studies that have demonstrated that resonance characteristics in male-to-female transsexuals can be changed to more closely approximate those of females through oral resonance therapy. This intervention study also demonstrates that a spontaneous increase in F0 is achieved during the course of therapy. Further, this study provides preliminary evidence to suggest that oral resonance therapy may be effective in increasing femininity of voice in male-to-female transsexual clients.

Key Words: Transsexual voice—Oral resonance—Formant frequencies—Fundamental frequency—Perceptual voice analysis.

INTRODUCTION

One of the greatest hurdles in the changeover of gender roles for many male-to-female transsexuals is that their voice, speech, and communication characteristics contrast with their desired gender. This discrepancy exposes the individual's biological gender, often resulting in a lack of social acceptance.[1,2] In these situations, a speech pathologist may be called upon to contribute to the gender reassignment process by facilitating change in communication behaviors that are incongruent with the desired gender role of the individual.

Accepted for publication May 9, 2006.

From the School of Human Communication Sciences, La Trobe University, Bundoora, Victoria, Australia.

Address correspondence and reprint requests to Georgia Dacakis, School of Human Communication Sciences, Level 3, Health Sciences Building 1, La Trobe University, Bundoora, Victoria, Australia 3086. E-mail: g.dacakis@latrobe.edu.au

Journal of Voice, Vol. 21, No. 5, pp. 591–603
0892-1997/$32.00

doi:10.1016/j.jvoice.2006.05.005

Voice therapy for male-to-female transsexuals traditionally focuses on increasing the pitch or mean fundamental frequency (F0) of the voice.[3–6] It has been established that increasing F0 results in an increase in the perception of femininity of voice in male-to-female transsexuals.[7] However, this increase is often not sufficient for the client to be perceived consistently as female.[8,9]

Altering oral resonance via lip spreading (retraction of the corners of the mouth) and the use of a more anterior tongue carriage during speech have been reported in the literature as potentially beneficial targets of therapy for male-to-female transsexuals[4,8,10,11]; however, there is little research evidence to support its use in intervention.

Previous studies have found that average vowel formant frequency values of males are approximately 20% lower than those of females.[12] Coleman[13] hypothesizes that this difference may be primarily due to the fact that females tend to have smaller physical dimensions of the cavities of the head and neck than men. Hence, the frequencies at which the vocal tract resonates in females can be expected to be higher. Gunzburger[10] proposes, however, that differences in vowel formant frequency values between men and women are too great to be due to anatomical differences alone. She suggests that speakers modify their vocal characteristics to conform to feminine and masculine speech stereotypes. One example of this may be that females use more lip spreading and a more anterior tongue carriage during speech than males, as these are considered characteristic of females. Gunzburger[10] suggests that it may be these speech differences which contribute to the differences in resonance and formant frequency values between males and females.

This theory is supported by research evidence demonstrating that formant values can be altered by changing dimensions of the vocal tract during speech. Changing the shape of the vocal tract at the mouth (eg, lip spreading) and changing where and to what extent the vocal tract is constricted (eg, tongue fronting) affect the length and shape of the vocal tract. This in turn affects the vowel formant frequency values produced. According to Ladefoged,[14] the frequency of the first formant is inversely related to vowel height. That is, the first formant frequency value increases as the tongue position lowers during the production of vowel sounds. The second formant frequency is known to be correlated to the degree of fronting of the tongue during speech; as the tongue moves forward in the mouth of the speaker, the second formant increases.[8,14] It has also been found that increased lip spreading during speech results in overall increased values of the third formant.[10,14]

Gunzburger[15] recruited six male-to-female transsexual clients who read a series of words in their ''female voice'' and then again in their ''male voice.'' Vowel formant frequency values for formant one (F1), formant two (F2), and formant three (F3) were obtained via spectrographic analysis. No significant differences were found between male and female modes for F1 or F2. F3, however, was found to have undergone a systematic upward shift in the central frequency in the ''female'' mode. Gunzburger suggested that this may have been due to shortening of the oral cavity length via lip spreading; however, this appeared to be speculation based on the premise that lip spreading during speech results in increased F3 values. No measures were taken to establish whether in fact the speakers were using more lip spreading in the female mode. Therefore, the differences between the modes of speaking which resulted in the increase in F3 frequencies are largely unknown.

In a study conducted by Gelfer and Schofield,[6] 15 male-to-female transsexual clients were recorded while reading ''The Rainbow''[16] passage and while producing two isolated vowels. Listeners were asked to identify each speaker as male or female and to rate the femininity-masculinity of the speaker's voice on a seven-point scale. Those clients perceived as female were found to have consistently higher vowel formant frequencies than those perceived as male. The second formant frequency of the vowel /i/ was moderately correlated with femininity scores, but not to a significant degree. However, of the 15 participants in the study, only three were perceived consistently as female. Due to the extremely small sample size of participants perceived as female, the generalizations that can be drawn from this study are limited.

Coleman's study[13] investigated how well listeners could identify speaker gender without

fundamental frequency cues. Ten male and 10 female subjects were recorded while reading "The Rainbow"[16] passage, and then while producing two vowels, using an electrolarynx with a fundamental frequency of 85 Hz. A listening panel made up of 15 students was able to identify speaker gender correctly 88% of the time. Listeners were also asked to indicate, on a seven-point scale, the confidence with which their selection of male versus female was made. Participants' vowel productions were later analyzed spectrographically to obtain formant frequency values (F1, F2, and F3). Coleman[13] calculated an average of the three formants to represent the overall vocal tract resonance of each subject (VTR score). VTR scores were found to be closely correlated with the degree of confidence with which listeners made their judgments. That is, listeners indicated more confidence in their judgments of gender for males with lower VTR scores and for females with higher VTR scores. This suggests that resonance characteristics of voice are powerful cues for identifying speaker gender.

Another study by Coleman[17] again made use of the electrolarynx, this time setting the F0 at 120 Hz to represent a male voice and 240 Hz to represent a female voice. Five males and five females were recorded while reading "The Rainbow"[16] passage with the electrolarynx at the two above-mentioned frequencies. Twenty-five listeners were asked to make a judgment of gender from a 5-second segment of each recording. Results indicated that even when the F0 was at a female level (240 Hz), listeners rated the voice as being male when the speaker was in fact male. This emphasizes that it is not sufficient to work solely on increasing F0 in male-to-female transsexuals, as this may not be enough for the individual to be perceived as female. Coleman attributes the results of this study to the resonance characteristics of the male speakers being an overriding factor in gender identification of voice. However, in both of his studies, extraneous speech cues such as rate, syllable stress, and articulation patterns were not controlled for and may have influenced the listeners' perceptions of gender. Despite these limitations, Coleman's results indicate that there is a relationship between vowel formant frequency values and perceptions of gender. This suggests that vocal resonance is likely to play an integral part in identifying speaker gender.

Mount and Salmon[8] reached similar conclusions during a longitudinal study of a male-to-female transsexual. After their client increased her F0 from 110 to 205 Hz, she was still not perceived as female over the telephone. After 11 months of therapy which included targeting forward tongue carriage, vowel formant frequency values had increased, with the greatest change being from 2092 to 2383 Hz for F2 on /i/. Once this increase in F2 had been achieved, the client was reported to be perceived as female over the telephone. This led Mount and Salmon[8] to remark that "vocal tract resonance characteristics may be the second most important acoustic cue to speaker identification" (p. 230). Although no correlations were performed to assess the relationship between F2 values and perception of femininity in this study, the results do indicate that altering tongue position may be effective in changing oral resonance and increasing listener perceptions of femininity of voice.

Although prior research in this area is limited, studies have indicated that formant frequency characteristics are important acoustic cues in gender identification.[17] Previous research also suggests that formant frequency values in transsexuals may be changed to approximate those of female speech by altering oral resonance via increasing forward tongue carriage and lip spreading.[8,10]

The aim of the present study was to investigate the effectiveness of oral resonance therapy on the perception of femininity of voice in a sample of 10 male-to-female transsexuals. Acoustic differences in vowel formant frequencies and F0 were also evaluated pre- and posttreatment to determine the effect of oral resonance therapy on the acoustic parameters of the voice. Additional aims were to determine client self-perceptions of the femininity of their voice pre- and posttherapy and their levels of satisfaction with their voice.

METHOD

Participants

Ten male-to-female transsexuals volunteered to participate in this study. None of the participants had undergone sexual reassignment surgery, and

none had commenced voice therapy to feminize their voice. Of the 10 participants, 6 were living full time in the female role. The remaining four participants were still living full time as males. Eight of the participants were enrolled in the Gender Dysphoria Program at Monash Medical Centre in Melbourne, Australia, while the remaining two were yet to enroll. The ages of the participants ranged from 25 to 64 years, with a mean of 40 years.

Participants were recruited from the waiting list for the Voice Clinic of the La Trobe University Communication Clinic in Melbourne, between March and August, 2002. Participants were contacted by mail, or by telephone if no address was supplied.

Therapy

Participants attended five therapy sessions targeting oral resonance. These sessions were scheduled at weekly intervals and were each of 45 minutes' duration.

A small number of references to alteration of oral resonance therapy in male-to-female transsexuals can be found in the literature. Comprehensive literature searching, however, did not reveal any details of intervention procedures for oral resonance therapy in this population. Therefore, for the purposes of this study, a specific therapy program was developed. Oral resonance techniques outlined by Boone[18] and Martin and Darnley[19] were used as a guide for the development of these procedures. To modify oral resonance, two therapy goals were developed; the first goal aimed to increase the use of lip spreading during speech for each participant; the second aimed to increase forward tongue carriage during speech for each participant. The two goals were instated separately and only targeted together when participants were able to use the techniques in complete sentences. Participants worked progressively through the following tasks:

- Auditory and visual discrimination of techniques from clinician-provided models (ie, discrimination between forward and backed tongue carriage and discrimination between use of lip spreading and lip rounding).
- Positive and negative practice of each technique (ie, contrasting the use of forward vs back tongue carriage and contrasting the use of lip spreading vs lip rounding).
- Practice of techniques in:
  (i) isolated vowels
  (ii) consonant-vowel and vowel-consonant syllables
  (iii) monosyllabic words
  (iv) bisyllabic words
  (v) phrases
  (vi) sentences (simultaneous use of both techniques)
  (vii) conversation (simultaneous use of both techniques).

Participants were asked to spend a minimum of 15–20 minutes per night practicing the use of the techniques while reading from homework sheets provided. At the beginning of each session, home practice was reviewed before progressing to the next stage of therapy.

Data collection

Audio recordings

Participants were audio-tape-recorded pre- and posttherapy while reading the first paragraph of "The Rainbow"[16] passage. Participants were given time to silently read over the passage prior to the recording, to ensure that unfamiliarity with the reading passage did not confound the results of the study. Participants were then instructed to use their everyday conversational speaking voice while reading the passage.

The recordings were made either in a quiet room with low ambient noise at the Speech Pathology Clinic at the Royal Victorian Eye and Ear Hospital in Melbourne, or in a sound-treated audiology suite at the La Trobe University Communication Clinic.

It would have been preferable to have used the same experimental setting for the audio recordings of all 10 participants; however, not all participants were willing or able to come to the one venue due to issues with traveling times, transportation, and parking expenses. The recordings of each participant pre- and posttreatment, however, were made in the same experimental setting.

The reading samples were recorded on a Sony digital audio tape (DAT) recorder, TCD-DIOPRO2 (Sony, Melbourne, Australia). Recording levels

were adjusted to prevent distortion of recordings. Recordings at the La Trobe University Communication Clinic were made using the Nakamichi CM-100 Electret Condenser Microphone (Nakamichi, Singapore) connected to a Sony DAT recorder. Recordings at the Royal Victorian Eye and Ear Hospital were made using the Shure SM48 unidirectional dynamic microphone (Shure, Niles, IL). The microphone used was set on a stand and placed 20 cm in front of the participant's mouth. To avoid unwanted ''pop'' (explosive breath sounds), the microphone was positioned 5 cm beneath and 5 cm to the left of the participant's mouth.

Ideally, the microphone used should have remained consistent for all participants involved in the study. This could not be arranged for practical reasons. It was however ensured that the same microphone was used both pre- and posttreatment for each participant.

Listener ratings

Twelve fourth-year speech pathology students were recruited to perceptually rate femininity of voice in the pre- and posttreatment recordings of each participant. Audio-taped recordings of each participant reading ''The Rainbow''[16] passage both pre- and posttreatment were randomized and dubbed onto one TDK DAT tape (TDK, Redhill, Surrey, UK). These recordings were then played back on the Sony DAT recorder using a Jazz amplified computer speaker, model J-215 (Jazz, Chung-Ho, Taiwan). Each student was given a series of visual analog scales (VASs), one for each recording. The students were instructed to indicate to what extent they perceived the voice as feminine in each recording by placing a vertical line through a 10-cm horizontal line, labeled ''very masculine'' at the left end and ''very feminine'' at the right end.

Participant self-ratings

The participants were asked to complete two VASs pre- and posttherapy. Participants were instructed to indicate on the first scale how masculine or feminine they felt their voice sounded by placing a vertical line through a 10-cm horizontal line, labeled ''very masculine'' at one end and ''very feminine'' at the other. Participants were then instructed to indicate on the second scale how satisfied they were with their voice, again by placing a vertical line through a 10-cm horizontal line. This time the labels on the extremes of the scale were ''very unsatisfied'' and ''very satisfied.''

VAS responses from both transsexual participants and student raters were scored by converting the distance from the left end of the scale, in centimeters, into a score from 0 to 10.

Listener intra- and interrater agreement

Intrarater agreement was evaluated by inserting four duplicate samples (two pre- and two posttreatment recordings) randomly onto the DAT tape containing the audio-taped recordings. Ratings given by individual judges were then compared to ensure that similar scores (within ± 1.5 cm) were given on both presentations of an identical recording.

Interrater agreement was evaluated by computing the median value for femininity scores given by raters. Each individual rater's score was then compared to this median score in order to determine whether the rater's score was consistent with those produced by the other raters. Scores within ± 1.5 cm of the median were considered to have adequate interrater agreement.

Acoustic analysis

The computer program *Computerized Speech Laboratory* (*CSL*, Model 4300B, Version 2.2, Kay Elemetrics Corporation, Lincoln Park, NJ) was used for acoustic analyses. Acoustic analysis of F0 was performed on each reading sample.

Further acoustic analysis consisted of extracting vowel formant frequencies from the vowels /a/, /i/, and /ʊ/ contained in the reading passages. These specific vowels were chosen as they represent extreme tongue and lip positions during speech production.

Traditionally, /u/ rather than /ʊ/ has been used by researchers when analyzing vowel formant frequencies as it is produced with more extreme tongue and lip positions (further back and more rounded, respectively). Also, as /u/ is a long vowel, it is generally easier to analyze spectrographically than the short vowel /ʊ/. However, as ''The Rainbow''[16] passage contained no suitable /u/ vowels for analysis, /ʊ/ was chosen as it is very similar in production to /u/, with a backed tongue position and lip rounding.

The readings of ''The Rainbow''[16] passage were transferred into *CSL* and edited to isolate, one-by-one, the vowels (/a/, /i/, and /ʊ/) from selected words within the passage. As connected speech samples were used (rather than isolated vowels), it is possible that there may have been some coarticulatory influences on the vowel formant frequency values used for analysis. To enhance the representativeness of the vowel formant frequencies obtained, it was ensured that the words from which the vowels were extracted were not spoken with stress or emphasis and were not preceded or followed by another vowel or nasal sound as this would be likely to influence the vowel formant frequency values. Furthermore, as the same reading passage was used pre- and posttreatment, any coarticulatory influences would be expected to be similar for both pre- and posttherapy measures.

CSL was used to produce a broadband spectrographic display to graphically plot out all vowel formants (0–6000 Hz) as a function of time and to produce a numerical display of these formant values. As these values are produced as a function of time, several formant frequency values were displayed for each vowel. To ensure consistency and to eliminate run-on effects from surrounding sounds, the vowel formant frequency values (F1, F2, and F3) were extracted from the midpoint of each vowel, at the most stationary portion of the formants.

As recommended by Ladefoged,[14] accuracy of formant frequencies obtained through *CSL* was ensured by manually estimating the formant frequency values through spectrographic analysis. As each vowel was isolated on the *CSL* display (prior to analysis), each formant band (F1, F2, and F3) was analyzed visually to estimate the value of each formant at the midpoint and most stationary portion of the formant. These results were then compared with the numerical results produced by *CSL* to ensure that no computer errors were made.

For each participant, the vowels (/a/, /i/, and /ʊ/) were extracted from two different words in the reading passage. The mean of the two vowel formant frequency values obtained was used to increase the reliability and representativeness of the formant frequency values.

RESULTS

All data were analyzed using the software package *SPSS* for Windows, Version 10.0.5 (SPSS Inc., Chicago, IL). During data screening, any outliers identified were eliminated from the data prior to further analysis. Data screening revealed that data for F0, as well as for all the individual formants of each vowel, were significantly skewed for pretreatment, posttreatment, or both. The ratio data collected (fundamental frequency and vowel formant frequency values) significantly violated the assumption of normality for the parametric *t* test. Therefore, its nonparametric equivalent, the Wilcoxon signed ranks test, was used for statistical analysis of all acoustic data collected. The Wilcoxon signed ranks test was also used for all perceptual data as this study used a repeated measures design, with all perceptual data collected on an ordinal scale.

Vowel formant frequency values

Group mean and median values of F1, F2, and F3 for all three vowels analyzed were higher posttreatment than pretreatment (Tables 1–3). The Wilcoxon

TABLE 1. *Group Means, Medians, Standard Deviations (SD), and Ranges for F1 During Production of /i/, /a/, and /ʊ/ Pre- and Posttreatment*

	N	Mean	Median	SD	Range
/i/ Pretreatment	10	323.6	329.8	53.7	229.5–380.5
/i/ Posttreatment	10	352.0	337.8	80.9	231.5–519.5
/a/ Pretreatment	10	748.1	740.3	79.6	626.0–875.0
/a/ Posttreatment	10	820.4	830.9	109.7	658.5–915.5
/ʊ/ Pretreatment	9	360.9	366.5	45.9	262.5–417.0
/ʊ/ Posttreatment	9	421.1	420.5	79.8	283.0–524.0

TABLE 2. *Group Means, Medians, Standard Deviations (SD), and Ranges for F2 During Production of /i/, /a/, and /ʊ/ Pre- and Posttreatment*

	N	Mean	Median	SD	Range
/i/ Pretreatment	9	2116.9	2043.5	162.1	1907.0–2400.0
/i/ Posttreatment	9	2192.1	2198.0	137.6	1953.0–2390.5
/a/ Pretreatment	9	1277.5	1369.5	123.7	1116.5–1398.0
/a/ Posttreatment	9	1459.7	1458.0	48.2	1372.5–1525.5
/ʊ/ Pretreatment	10	981.5	983.3	74.7	880.0–1100.0
/ʊ/ Posttreatment	10	1037.7	1023.8	143.4	788.5–1280.5

signed ranks test showed that vowel formant frequency values were significantly higher posttreatment for F1 values of /a/ and /ʊ/ ($T = 5$, $Z = 2.295$, $P = 0.022$ and $T = 3$, $Z = 2.310$, $P = 0.021$, respectively), F2 values of /a/ ($T = 1$, $Z = 2.547$, $P = 0.011$), and F3 values for all three vowels, /i/, /a/, and /ʊ/ ($T = 1$, $Z = 2.701$, $P = 0.007$; $T = 0$, $Z = 2.668$, $P = 0.008$; and $T = 0$, $Z = 2.668$, $P = 0.008$, respectively).

Mean fundamental frequency (F0)

The Wilcoxon signed ranks test demonstrated that F0 values were significantly higher posttreatment than pretreatment ($T = 3$, $Z = 2.497$, $P = 0.013$). Table 4 displays the means, medians, standard deviations, and ranges for participants' F0 of voice pre- and posttreatment.

Listener ratings of femininity of voice

Eight raters out of the 12 produced ratings that scored within ±1.5 cm of each other on the VAS, for all duplicate recordings. The ratings of the four raters whose scores were more than ±1.5 cm apart for duplicate recordings were considered to indicate unsatisfactory intrarater agreement, and so were eliminated from the results prior to assessing interrater agreement.

Only 7 out of the 10 transsexual participants (70%) were rated within ±1.5 cm of the median by all eight raters for their pretreatment recordings. Posttreatment ratings showed that only 6 out of the 10 participants (60%) were rated within ±1.5 cm of the median by all eight raters. Interrater agreement of listener ratings was therefore deemed too low to conduct inferential statistics; however, Figures 1–10 show the scores given by the eight raters for each participant pre- and posttreatment for visual analysis of trends in the data. Readers should note that the numerical scale on the *y*-axis has been varied in Figures 1–10 to allow for clearer visual representation of changes pre- and posttherapy for individual participants.

It is difficult to deduce an overall trend from the ratings given to all 10 participants. Participants 2, 6, 7, and 9 showed obvious increases in perceived femininity of voice following treatment. Participants 1, 5, and 8 appear to have general, albeit variable, increases in femininity following treatment. Participant 10, however, does not appear to show any obvious change in either direction, whereas

TABLE 3. *Group Means, Medians, Standard Deviations (SD), and Ranges for F3 During Production of /i/, /a/, and /ʊ/ Pre- and Posttreatment*

	N	Mean	Median	SD	Range
/i/ Pretreatment	9	2551.6	2557.5	168.2	2246.5–2771.0
/i/ Posttreatment	9	2805.5	2776.0	172.8	2544.5–3046.0
/a/ Pretreatment	10	2508.7	2487.5	145.0	2332.5–2789.0
/a/ Posttreatment	10	2732.0	2696.5	155.4	2516.0–2968.5
/ʊ/ Pretreatment	10	2508.7	2487.5	145.0	2332.5–2789.0
/ʊ/ Posttreatment	10	2732.0	2696.5	155.4	2516.0–2968.5

TABLE 4. *Group Means, Medians, Standard Deviations (SD), and Ranges for Mean Fundamental Frequency Pre- and Posttreatment*

	N	Mean	Median	SD	Range
Pretreatment	10	119.4	115.2	16.1	100.0–152.8
Posttreatment	10	133.3	129.3	20.4	101.7–172.5

participants 3 and 4 show general decreases in perception of femininity following treatment.

Participant self-ratings

The Wilcoxon signed ranks test demonstrated that self-ratings of femininity of voice were significantly higher posttreatment than pretreatment ($T = 0$, $Z = 2.803$, $P = 0.005$), as were self-ratings of satisfaction with voice ($T = 8$, $Z = 1.988$, $P = 0.047$). Tables 5 and 6 display the means, medians, standard deviations, and ranges for participant self-ratings of femininity and satisfaction with voice, respectively.

DISCUSSION

The present study investigated the effectiveness of oral resonance therapy on the perception of femininity of voice in male-to-female transsexuals. The findings provide preliminary evidence to suggest that oral resonance therapy may be effective in changing the resonance characteristics in male-to-female transsexual clients, and that this, in turn, may increase the perception of femininity in these clients.

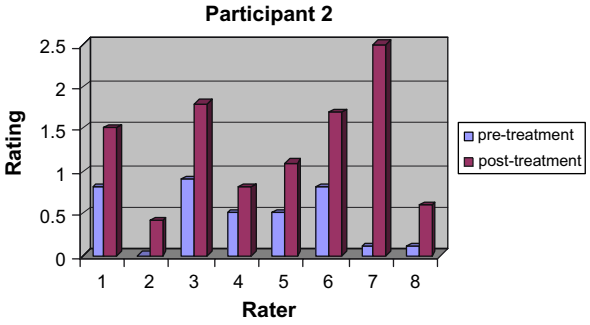

FIGURE 2. Listener ratings of perception of femininity of voice pre- and posttreatment for participant 2, where a rating of 0 indicates "very masculine" and a rating of 10 indicates "very feminine."

Vowel formant frequencies

The present study demonstrated that following treatment there was a general increase in vowel formant frequency values (F1, F2, and F3) for all vowels (/i/, /a/, and /ʊ/) for the 10 transsexual participants. This increase was statistically significant for F1 values of /a/ and /ʊ/, F2 values of /a/, and F3 values for all three vowels.

As discussed earlier, F2 is known to be related to the degree of backness or forwardness of tongue carriage[8,14]; as forward tongue carriage increases, F2 values also increase. Thus, the results suggest that the treatment administered was successful in achieving the therapy goal of increasing forward tongue carriage during speech in the participants

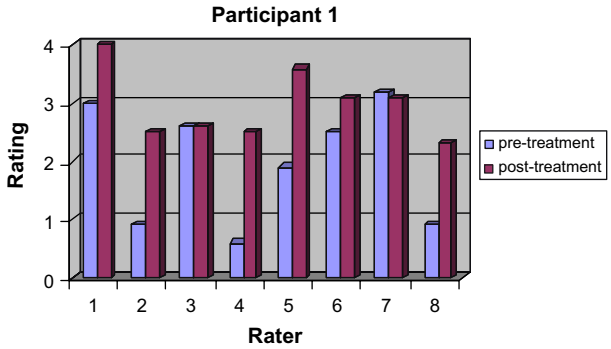

FIGURE 1. Listener ratings of perception of femininity of voice pre- and posttreatment for participant 1, where a rating of 0 indicates "very masculine" and a rating of 10 indicates "very feminine."

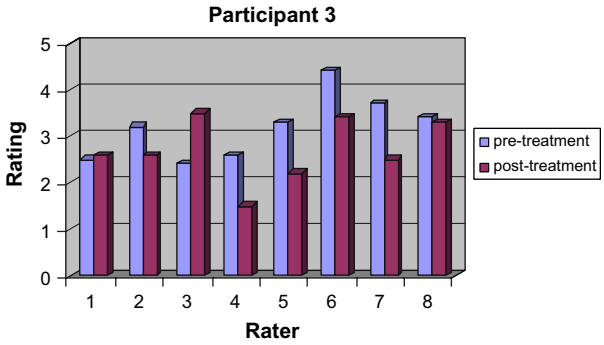

FIGURE 3. Listener ratings of perception of femininity of voice pre- and posttreatment for participant 3, where a rating of 0 indicates "very masculine" and a rating of 10 indicates "very feminine."

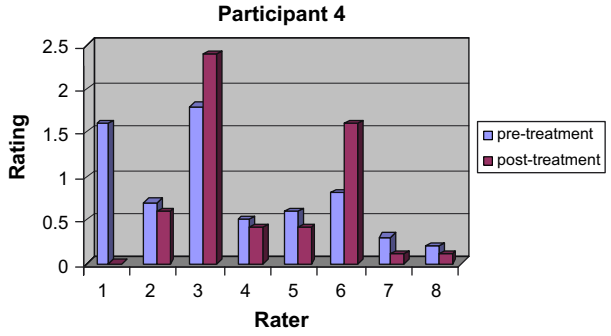

FIGURE 4. Listener ratings of perception of femininity of voice pre- and posttreatment for participant 4, where a rating of 0 indicates "very masculine" and a rating of 10 indicates "very feminine."

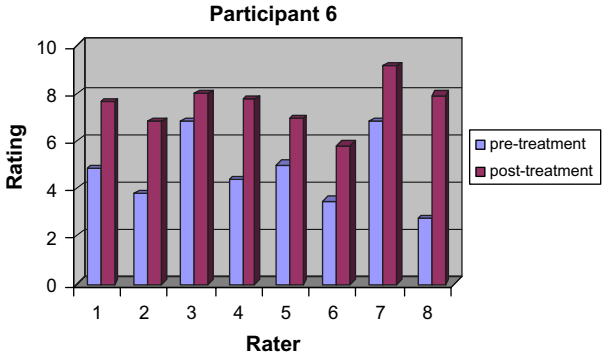

FIGURE 6. Listener ratings of perception of femininity of voice pre- and posttreatment for participant 6, where a rating of 0 indicates "very masculine" and a rating of 10 indicates "very feminine."

of the study. A significant increase was achieved on the vowel /a/, which is the most backed vowel in Australian English.[14] This may be attributable to the fact that a backed vowel leaves more scope for increases in forward tongue carriage than a front vowel such as /i/.

Mount and Salmon[8] also targeted forward tongue carriage in a single case study with one male-to-female transsexual client. They reported that it took 11 months of weekly therapy sessions before a significant change in F2 values was established. Therefore, it may be that the short period of therapy in the present study was inadequate to achieve significant changes in such a habitual behavior as tongue carriage during speech, for all vowels measured (/i/, /a/, and /ʊ/). It is possible that if a longer block of therapy were to be provided to these participants, greater increases in F2 values would be achieved for all vowels.

F3 values are known to be related to the degree of lip spreading used by speakers.[14] The significant increase in F3 values for all these vowels posttreatment suggests that the treatment successfully achieved the therapy goal of increasing lip spreading during speech.

The technique of lip spreading appeared to be easily and quickly mastered by the transsexual participants in only five therapy sessions. If it can be established that increasing lip spreading in male-to-female transsexual clients results in, or contributes to, an increase in the perception of femininity of voice, then this goal should be readily incorporated into therapy in this population.

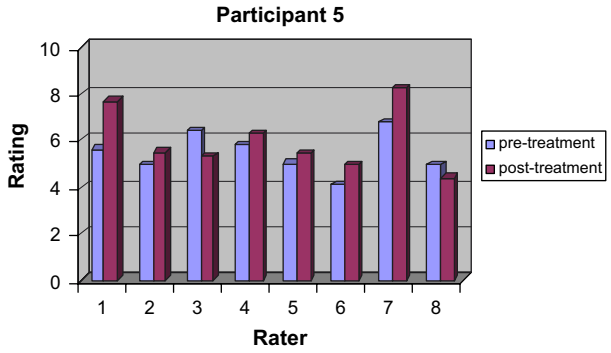

FIGURE 5. Listener ratings of perception of femininity of voice pre- and posttreatment for participant 5, where a rating of 0 indicates "very masculine" and a rating of 10 indicates "very feminine."

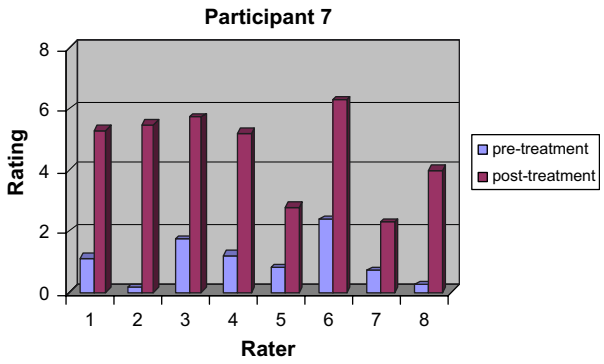

FIGURE 7. Listener ratings of perception of femininity of voice pre- and posttreatment for participant 7, where a rating of 0 indicates "very masculine" and a rating of 10 indicates "very feminine."

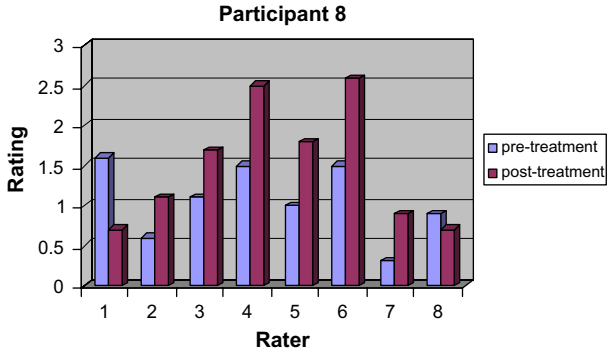

FIGURE 8. Listener ratings of perception of femininity of voice pre- and posttreatment for participant 8, where a rating of 0 indicates ''very masculine'' and a rating of 10 indicates ''very feminine.''

F1 is known to be inversely related to vowel height[14]; that is, the lower the tongue is held in the mouth during speech, the higher the F1 values generated. Therefore, the general increase in F1 values suggests that participants lowered their habitual tongue height during speech as a result of therapy. During introduction of the technique of forward tongue carriage, participants were instructed to rest their tongue flat on the floor of their mouth, with the tongue tip extending to touch the back of the bottom front teeth. Participants were instructed to use this tactile feedback to self-monitor their forward tongue carriage during sessions. It is therefore possible that this instruction resulted in participants lowering their habitual tongue height during speech, resulting in the overall increase in F1 values.

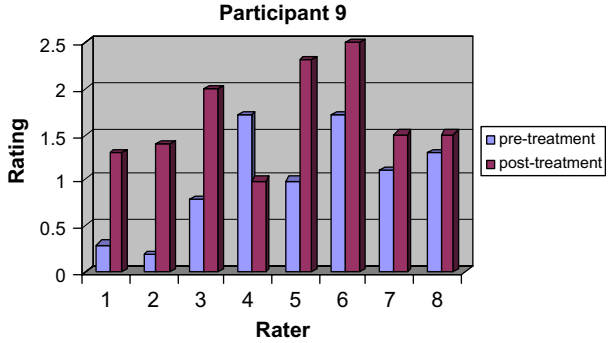

FIGURE 9. Listener ratings of perception of femininity of voice pre- and posttreatment for participant 9, where a rating of 0 indicates ''very masculine'' and a rating of 10 indicates ''very feminine.''

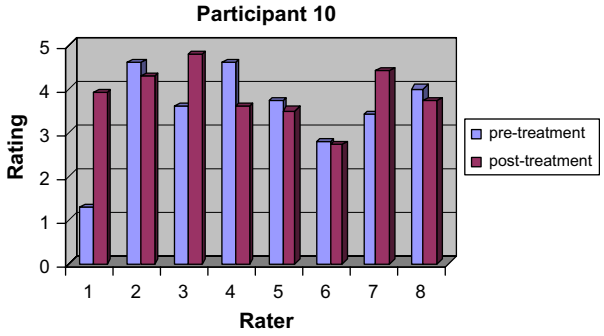

FIGURE 10. Listener ratings of perception of femininity of voice pre- and posttreatment for participant 10, where a rating of 0 indicates ''very masculine''and a rating of 10 indicates ''very feminine.''

Mean fundamental frequency (F0)

The results of the present study demonstrate that a significant increase in F0 of up to 30 Hz was achieved by participants following treatment, even though F0 was not specifically targeted as a goal of this treatment study. A number of hypotheses for this observed increase in F0 can be posited.

The commencement of therapy may have resulted in increased levels of motivation in participants to alter their voice to be more consistent with that of female speakers. As higher F0 is associated with perceptions of femininity,[3,4] it is probable that participants consciously and independently targeted an increase in their F0. This may also have been related to participants becoming more aware of the incongruence between their voice and gender after commencing therapy.

Further, the resultant increase in F0 could simply have arisen as a side effect of participants modeling the voice of a female clinician during intervention sessions. As a substantial amount of the therapy administered during tasks involved repetition of a clinician model, it is possible that these imitative tasks resulted in participants increasing their F0 to more closely resemble that of the clinician.

TABLE 5. *Group Means, Medians, Standard Deviations (SD), and Ranges for Participant Self-ratings of Femininity of Voice Pre- and Posttreatment*

	N	Mean	Median	SD	Range
Pretreatment	10	3.4	3.1	1.6	0.7–5.8
Posttreatment	10	5.6	5.7	1.2	3.6–7.3

TABLE 6. *Group Means, Medians, Standard Deviations (SD), and Ranges for Participant Self-ratings of Satisfaction of Voice Pre- and Posttreatment*

	N	Mean	Median	SD	Range
Pretreatment	10	3.4	2.7	2.2	0.5–7.7
Posttreatment	10	5.4	5.4	1.1	3.6–7.1

According to Oates and Dacakis,[4] the range of F0 for females is between 145 and 275 Hz, whereas the range of F0 for males is between 80 and 165 Hz. Therefore, for some male-to-female transsexual clients it is possible that an increase in F0 of 30 Hz would be sufficient to increase F0 into the female range. The results of this study imply that specifically targeting an increase in pitch may not be necessary for some male-to-female transsexual clients, because an increase of 30 Hz may occur incidentally while targeting other aspects of voice, such as oral resonance.

Listener ratings—interrater agreement

As discussed above, the interrater agreement for listener ratings in this study was deemed too poor to conduct inferential statistics. The use of a VAS, rather than an equal-appearing-interval scale (EAI) with five to seven points, may have contributed to poor interrater reliability in this study. The VAS was used for listener ratings of femininity, rather than an ordinal scale such as an EAI, due to claims that the VAS is more sensitive in measuring change.[20] Due to the lack of previous research in the area of oral resonance therapy with male-to-female transsexuals, the degree or direction of expected change in perceptions of femininity of voice was not known. Therefore, the VAS was used to increase the chances of finding changes in perception of femininity of voice pre- and posttreatment, even if those changes were only slight. Ratings of femininity of voice, however, may be best suited to an EAI with five to seven points, where additional guidance may be provided to the rater by the five to seven marker points along the scale. However, it is unlikely that the use of the VAS would be the sole cause for such poor levels of agreement.

It appears that listener perceptions of gender and femininity/masculinity of voice vary considerably between individuals. This can be seen in previous studies such as Coleman,[13] where listeners differed in their perceptions of gender as measured with dichotomous ratings; that is, some listeners judged a voice as male, whereas other listeners judged the same voice as female. Listener perceptions may be influenced by a multitude of factors. It is possible that life experiences such as exposure to different cultures and nationalities and even exposure to speakers with varying vocal qualities may influence a listener's perception of gender of voice. These factors may have contributed to the poor interrater agreement observed in the present study.

The poor levels of interrater agreement of listener judgments of femininity of voice for the present study have further implications for future research in this area. To definitively evaluate therapy targeting an increase in femininity of voice, reliable listener ratings must be obtained. Therefore, the issue of poor interrater reliability needs to be addressed before further studies can be conducted.

Future studies may be well advised to use speech pathologists who are experienced in the area of voice to carry out perceptual ratings. As the auditory-perceptual rating of femininity of voice on a scale (whether it be a VAS or an EAI) is quite an unfamiliar task for most people, better agreement might be obtained by using listeners who are experienced in the area of critically evaluating voice.

Trends observed from listener perceptions

Unfortunately, the poor interrater reliability achieved in the present study prevented inferential analysis of listener ratings of femininity from pre- to posttherapy. Visual analysis of these ratings, however, has provided some preliminary information regarding the effectiveness of oral resonance therapy on the perception of femininity of voice in male-to-female transsexuals.

The majority of participants (70%) achieved general increases in the perception of femininity of voice following treatment, as judged by the eight speech pathology student raters. Because increases both in vowel formant frequency values and in F0 were achieved as a result of this study, it is likely

that changes in oral resonance as well as increases in F0 contributed to the increase in perception of femininity of voice.

Ideally, correlations would be conducted on the data to confirm the relationship between increases in vowel formant frequency values and perceptions of femininity of voice, and also the relationship between F0 and perceptions of femininity of voice. Due to the poor interrater agreement achieved in this study, however, no correlations could be performed. Therefore, further research is warranted to both verify the preliminary results achieved in the present study and investigate the basis for any increases in femininity of voice.

Given that an increase in F0 was achieved without directly targeting F0 in therapy, it is possible that other aspects of speech that are known to be correlated with gender such as rate, intonation, syllable stress, and articulation patterns may have also changed during the course of therapy and contributed to perceptions of femininity. Because this was a small pilot study, however, these components of speech were not able to be systematically investigated.

Participant self-ratings

Participant self-ratings of both satisfaction and femininity of voice were found to be significantly higher posttreatment than pretreatment. In fact, all 10 participants of the study perceived themselves as sounding more feminine following therapy, whereas listeners only perceived 7 of these participants as sounding more feminine following therapy.

The validity of participant self-ratings of voice can, however, be questioned. It is likely that participants commenced therapy with the expectation that their femininity of voice would increase. This preconceived expectation may have influenced participant ratings. Further, as participants were informed that they were participating in a research project that aimed to increase the femininity of voice, participants may have indicated responses that they believed would please the researcher.

CONCLUSION

The present study demonstrated that forward tongue carriage and lip spreading can be increased to increase the vowel formant frequency values of male-to-female transsexual clients. This therapy may also achieve a simultaneous increase in F0 in this client population.

Despite the lack of inferential analysis of changes in listener perceptions of femininity of voice, trends in the data suggest that the oral resonance therapy administered during this study was associated with increased listener perceptions of femininity of voice for the majority of participants. That is, listeners perceived the participants of the study to sound more feminine following treatment. Participant self-ratings of femininity of voice also support this finding.

The preliminary evidence obtained from the present study suggests that oral resonance therapy may be effective in increasing the perception of femininity of voice in male-to-female transsexuals. However, further research is warranted to build on the preliminary evidence provided by this study. It will also be important to establish the relative contributions of oral resonance changes and changes in F0 to perceptions of femininity of voice.

Acknowledgments: We would like to convey our thanks to the participants of the study for all their commitment and hard work. We also thank Sheryl Mailing and Anne Yorston (Royal Victorian Eye and Ear Hospital) for their support and assistance throughout this project.

REFERENCES

1. Colton RH, Casper JK. *Understanding Voice Problems. A Physiological Perspective for Diagnosis and Treatment*. 2nd ed. Baltimore, MD: Lippincott Williams and Wilkins; 1996.
2. Challoner J. The voice of the transsexual. In: Fawcus M, ed. *Voice Disorders and Their Management*. London: Chapman and Hall; 1991:224–239.
3. Dacakis G. Long-term maintenance of fundamental frequency increases in male-to-female transsexuals. *J Voice*. 2000;14:549–556.
4. Oates J, Dacakis G. Voice change in transsexuals. *Venerology.* 1997;10:178–187.
5. Kalra MA. Voice therapy with a transsexual. Paper presented at: the American Speech and Hearing Association Convention; 1977; Chicago, IL.
6. Gelfer MP, Schofield KJ. Comparison of acoustic and perceptual measures of voice in male-to-female transsexuals perceived as female versus those perceived as male. *J Voice*. 2000;14:22–33.
7. Spencer LE. Speech characteristics of male-to-female transsexuals: a perceptual and acoustic study. *Folia Phoniatr.* 1988;40:31–42.

8. Mount MH, Salmon SJ. Changing the vocal characteristics of a post-operative transsexual patient: a longitudinal study. *J Commun Disord*. 1988;21:229–238.
9. Bralley RC, Bull GL, Gore CH, Edgerton MT. Evaluation of vocal pitch in male transsexuals. *J Commun Disord*. 1978;11:443–449.
10. Gunzburger D. Acoustic and perceptual implications of the transsexual voice. *Arch Sex Behav*. 1995;24:339–348.
11. de Bruin MD, Coerts MJ, Greven AJ. Speech therapy in the management of male-to-female transsexuals. *Folia Phoniatr*. 2000;52:220–227.
12. Coleman RO. Acoustic correlates of speaker sex identification: implications for the transsexual voice. *J Sex Res*. 1983;19:293–295.
13. Coleman RO. Male and female voice quality and its relationship to vowel formant frequencies. *J Speech Hear Res*. 1971;14:565–577.
14. Ladefoged P. *A Course in Phonetics*. Orlando, FL: Harcourt Brace Company; 1993.
15. Gunzburger D. Voice adaptation by transsexuals. *Clin Linguist Phon*. 1989;3:163–172.
16. Fairbanks GE. *Voice and Articulation Drillbook*. 2nd ed. New York, NY: Harper; 1960.
17. Coleman RO. A comparison of the contributions of two voice quality characteristics to the perception of maleness and femaleness in the voice. *J Speech Hear Res*. 1976;19:168–180.
18. Boone D. *The Voice and Voice Therapy*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall Inc.; 1977.
19. Martin S, Darnley L. *The Voice Sourcebook*. Oxon: Winslow Press; 1992.
20. McDowell I, Newell C. *Measuring Health: A Guide to Rating Scales and Questionnaires*. 2nd ed. New York, NY: Oxford University Press; 1996.